

PARTH KATLANA

katlanaparth@gmail.com | (984) 286-6759 | Santa Clara, CA | [LinkedIn](#) | [Github](#) | [Portfolio](#)

SUMMARY

- Data Scientist / ML Engineer with 3+ years building and deploying production AI systems from **agentic LLM workflows and RAG pipelines** to distributed deep learning on cloud-native infrastructure.
- Full ML lifecycle: data engineering, distributed model training, **experimentation (A/B testing)**, MLOps, and production deployment of scalable, reliable systems.
- Experienced in translating complex analyses into business impact through stakeholder communication and statistical rigor, **co-author of 3 papers on LLM reliability and evaluation**.

WORK EXPERIENCE

Data Scientist, NCICS - NOAA Affiliate | Remote

Jan '25 - Present

- Secured **\$500k NASA grant** by designing a novel forecasting prototype that reduced RMSE by 17.4% vs the CNN baselines.
- Designed and shipped a **production agentic AI system** (LangChain, RAG, LLMs) that autonomously plans multi-step workflows, retrieving context, calling tools, running analyses, generating structured reports, over 12–25 TB/mo of satellite data, cutting time-to-insight from ~3 hrs to 20 min.
- Built an **evolutionary agentic system** where LLM agents propose model-architecture mutations, orchestrate training & evaluation via Step Functions, and select optimal designs through multi-objective fitness with built-in failure handling.
- **Productionized deep learning models with FSDP and CUDA** distributed training, reducing GPU training cost by 40% and tile inference latency to 250 ms through parallelization.
- Designed and deployed **end-to-end ML pipelines** on AWS (Databricks, SageMaker, MLflow) for distributed preprocessing of geospatial and time-series data, with CI/CD automation, reducing deployment time by 30%.
- Engineered **A/B testing** with CDC stakeholders to evaluate indicator configurations, identifying an optimal signal-to-noise ratio that reduced false positive alerts by 12.5%.

Data Scientist, Laboratory of Analytical Sciences | Raleigh, NC

May '23 – Aug '24

- Orchestrated an **LLM ensemble (Mistral, Llama-2, GPT-4)** with a ranking layer that transforms unstructured text into structured, data-rich reports, reducing review time by 50%.
- Built an **LLM evaluation framework** with a metric-ensemble **hallucination detector** (0.91 Pearson correlation with human judgment) to validate factual consistency, fairness, and robustness, **published (2 papers)**.
- Fine-tuned in-house **CLIP and YOLOv8** models for domain-specific detection and retrieval, increasing mAP50-95 by 15%.
- Developed and tuned classification models (**XGBoost, Logistic Regression**) with Bayesian optimization to improve predictive accuracy for risk assessment.
- Led statistical and predictive modeling across the ML lifecycle (Python, SQL, Spark), driving a 15% increase in conversion through data-driven insights; authored documentation and presented to leadership, influencing roadmap decisions.

Data Scientist, Kion Technologies | India

Aug '19 – Jan '22

- Scaled an **A/B testing infrastructure** to 10M+ daily active users, reducing system latency by 40% through ML model integration and performance tuning.
- Performed **credit risk profiling and customer segmentation** (logistic regression, K-Means) on structured and behavioral metadata, refining risk assessment strategies.
- Optimized **ETL pipelines** across Snowflake and SQL Server, building regression and multivariate models in Python, enabling real-time KPI visibility and cutting down operation costs by 15%.
- Built data visualizations and reports (Power BI, Tableau, Matplotlib) that improved interpretability and increased marketing campaign effectiveness by 25%.

TECHNICAL SKILLS

Languages & Databases: Python, SQL, R, Java, PostgreSQL, MySQL, MongoDB, NoSQL, Snowflake

GenAI & LLMs: LangChain, RAG, Agentic Workflows, LLM Evaluation, Prompt Engineering, Fine-tuning, Vector Db, Hugging Face

Machine Learning & Deep Learning: PyTorch, TensorFlow, Scikit-learn, XGBoost, Transformers, BERT, CNNs, CLIP, YOLOv8, Bayesian Optimization, Distributed Training (FSDP/CUDA), Unsupervised Learning, K-means clustering, Regression

Cloud & MLOps: AWS (S3, SageMaker, Bedrock, Step Functions), GCP, Databricks, Spark, Kafka, Airflow, MLflow, Docker.

Analytics & Tools: Databricks, Tableau, Power BI, Matplotlib, Git, JIRA, CI/CD, Cursor, A/B Testing, ETL. Hypothesis testing

PROJECT

Agentic Credit-Risk Underwriting - Multi-agent system (LangGraph, LLMs) automating PD scoring (XGBoost), RAG-based Basel III/IV regulatory validation, and SHAP-driven decision explanations.

EDUCATION

North Carolina State University, Raleigh, NC | *Master of Computer Science*

May 2024

Symbiosis University, India | *Bachelor of Engineering, Computer Science*

Jul 2022

ACADEMIC PUBLICATIONS

- Parth K., Grant Forbes “Metric Ensemble for hallucination detection”. [Link to Paper](#)
- Zeydy Ortiz, Parth K. “Uncovering Factual Consistency Errors”. [Link to Paper](#)
- Marley C., Parth K., Chris K., Akhil N. “General Interface for Factored State Optimization Problems”. [Link to Paper](#)